Electronic Colloquium on Computational Complexity, Revision 4 of Report No. 130 (2026)

Optimal Bitwise Pseudorandom Generators for Modular Sums Modulo Every Fixed Integer

Gonen Krak[†]

Abstract

We study pseudorandom bit generators for Boolean linear sums modulo a fixed integer M. A source $X \in \{0,1\}^n$ succeeds with error ε if, for every $a \in \mathbb{Z}_M^n$, the residue

$$\sum_{i=1}^{n} a_i X_i \pmod M$$

is ε-close in statistical distance to the corresponding residue under independent uniform bits. Lovett, Reingold, Trevisan, and Vadhan obtained the optimal seed length for prime-power moduli. We prove the same bound for every fixed modulus:

$$O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

The proof builds the generator in stages. It keeps two quantities separate: the prime field $\mathbb{F}_p$ containing the generator's output, and the modulus m used by the test. A state (p, m) means that we have a PRG on $\mathbb{F}_p^n$ that fools every modulus-m phase whose coordinates contribute separately.

We use two steps to move between states. The first step, LIFT_p, changes the modulus from m to pm. The second step, $\mathrm{TRANSFER}_{r\to p}$, changes the alphabet from $\mathbb{F}_r$ to a smaller field $\mathbb{F}_p$.

Starting from the trivial modulus 1, we process the prime factors of M until we reach the Boolean state $(2, M)$.

[†]Tel Aviv University. gonenok10@gmail.com

ISSN 1433-8092

Contents

1 Introduction

A distribution over $\{0,1\}^n$ is *pseudorandom* for a class of tests if no test in the class can distinguish it from n independent uniform bits. The smaller the class, the shorter the seed one may hope for, and for the simplest nontrivial class — linear tests modulo 2 — the problem was solved in the seminal work of Naor and Naor [NN93]: *small-bias generators* fool all parities with error ε using a seed of length $O(\log n + \log \frac{1}{\varepsilon})$, which is optimal up to constant factors [AGHP92]. Small-bias generators have since become a basic primitive of pseudorandomness, and a natural next question is what happens when the modulus 2 is replaced by an arbitrary fixed modulus M.

Concretely, for a coefficient vector $a \in \mathbb{Z}_M^n$ consider the modular sum

$$S_a^{(M)}(x) = \sum_{i=1}^{n} a_i x_i \pmod{M}, \qquad x \in \{0,1\}^n.$$

A bit distribution X *ε-fools Boolean linear sums modulo M* if, for every $a \in \mathbb{Z}_M^n$, the distribution of $S_a^{(M)}(X)$ is ε-close in statistical distance to that of $S_a^{(M)}(U_n)$. Modular sums are among the simplest tests computable in logarithmic space; they are exactly the read-once products over the cyclic group $\mathbb{Z}_M$, that is, tests of the form $\prod_i g_i^{x_i}$ with $g_i \in \mathbb{Z}_M$; and fooling them is a necessary step in several programs on pseudorandomness for group products and read-once branching programs.

Lovett, Reingold, Trevisan, and Vadhan [LRTV09] constructed generators with the optimal seed length $O_M(\log n + \log \frac{1}{\varepsilon})$ for every fixed *prime-power* modulus $M = p^k$. Their second construction works for every M, at seed length

$$O\left(\log n + \log \frac{M}{\varepsilon} \cdot \log\left(M \log \frac{1}{\varepsilon}\right)\right),$$

which for fixed M is $O_M\big(\log n + \log \frac{1}{\varepsilon} \cdot \log\log \frac{1}{\varepsilon}\big)$: optimal in n, but short of optimal by a $\log\log \frac{1}{\varepsilon}$ factor. In the years since, the composite case has resisted matching the prime-power one. The recent work of Lee and Viola [LV25], which achieves optimal seed length $c_G \log(n/\varepsilon)$ for read-once group products over every p-group, explicitly identifies the handling of composite moduli as a challenge reminiscent of a known roadblock in circuit complexity. In particular, prior to this work no generator with optimal seed length was known even for $M = 6$, the smallest modulus with two distinct prime factors.

1.1 Our result

We remove the restriction to prime powers entirely.

Theorem 1.1 (Main result; informal)**.** *For every fixed integer $M \geq 2$ there is an explicit generator that ε-fools all Boolean linear sums modulo M with seed length*

$$O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

The formal statement is Theorem 7.3. The seed length is optimal up to the constant depending on M: any generator fooling linear sums modulo M also fools them modulo any prime divisor of M (Proposition 7.4), where the bound $\Omega(\log n + \log \frac{1}{\varepsilon})$ is standard.

The generator is built from off-the-shelf components. The only pseudorandomness primitive imported is explicit small bias over fixed prime fields [NN93], [AMN98]; everything else, including the base case, is proved here. What the construction does with those components is where the content lies.

AI-use disclosure. The exposition and parameter calculations were reorganized and cross-checked with the assistance of ChatGPT and Claude; the author retains responsibility for every statement and proof.

1.2 Related work

Modular sums. Naor and Naor [NN93] solved the case $M = 2$ optimally. Lovett, Reingold, Trevisan, and Vadhan [LRTV09] gave the two constructions displayed above. Gopalan, Kane, and Meka [GKM18] took a different route, constructing a single generator that fools the discrete Fourier transforms of linear functions and thereby all modular tests with modulus up to M simultaneously, with seed length

$$O\left(\log \frac{Mn}{\varepsilon} \cdot \left(\log\log \frac{Mn}{\varepsilon}\right)^2\right).$$

The strength of their result is its uniformity: one generator for every modulus up to M, with near-logarithmic dependence on M. For a fixed modulus it does not improve on [LRTV09]. What Theorem 1.1 does is remove the residual $\log\log \frac{1}{\varepsilon}$ loss of the LRTV arbitrary-modulus construction, for every fixed modulus rather than for prime powers alone.

Group products and branching programs. Linear sums modulo M are read-once products over the cyclic group $\mathbb{Z}_M$, a commutative special case of read-once permutation branching programs. For group products, Meka and Zuckerman [MZ09] gave generators over S_3 with polynomially small error at seed length $O(\log n)$, handling the modulus-3 part through generators for low-degree polynomials over $\mathbb{F}_3$. Lee and Viola [LV25] recently achieved optimal seed length $c_G \log(n/\varepsilon)$ for read-once group products over every p-group, and nearly optimal seed length over every commutative group; they point to composite moduli as the boundary of current techniques. Theorem 1.1 gives the exactly optimal seed length for every cyclic group $\mathbb{Z}_M$, including those that are not p-groups. By a standard character argument it should also yield the corresponding statement for every finite abelian group, with the error adjusted by a constant depending on the group; we do not pursue that here.

The Meka–Zuckerman conjecture. For $M = 6$, the final generator has a familiar form: it is the XOR of a small-bias distribution and a bit distribution that fools all linear sums modulo 3. This is exactly the type of generator that Meka and Zuckerman conjectured would fool width-3 read-once permutation branching programs [MZ09].

Our theorem does not prove that conjecture. Linear tests modulo 6 correspond to

$$\mathbb{Z}_6 \cong \mathbb{Z}_3 \times \mathbb{Z}_2,$$

whereas width-3 permutation programs involve

$$S_3 \cong \mathbb{Z}_3 \rtimes \mathbb{Z}_2.$$

In S_3, the modulo-2 part can reverse the modulo-3 part, creating a dependence on earlier input bits that our analysis does not handle.

Relation to degree descent over $\mathbb{F}_2$. Our technique resembles a line of work on pseudorandom generators for low-degree polynomials [BV10, Lov09, Vio09]. In those works, Cauchy–Schwarz introduces a directional derivative, each derivative lowers the polynomial degree, and the induction ends at a small-bias base case. Lemma 4.4 uses the same analytic pattern. The algebraic step in this paper also lowers degree, but only after changing the alphabet to the prime field $\mathbb{F}_p$ and reducing one local exponent modulo p. Every function on $\mathbb{F}_p$ then has degree at most $p-1$, so $p-1$ finite differences make it constant modulo p and reveal one factor p in the modulus. The overall construction therefore descends in the modulus while the inner algebra descends in degree. The need to change the alphabet is the new feature: no analogous bounded-degree representation over the Boolean field is available for counting modulo an odd prime, as reflected by the Razborov–Smolensky method [Smo87].

Organization. Section 2 gives the complete recursive route before any technical machinery is introduced. Section 3 defines the intermediate test classes and the error measure. Section 4 proves LIFT using polynomial degree reduction and a one-step derivative-lifting lemma. Section 5 proves $\mathrm{TRANSFER}$ by combining a small-bias estimate for low-weight tests with a cancellation estimate for high-weight tests. Section 6 follows the prime factors of M, and Section 7 converts the final phase guarantee into statistical distance and proves optimality.

2 The recursive construction plan

The main theorem concerns a Boolean generator for modulus M. Before defining phases, derivatives, or Fourier coefficients, this section gives the entire route of the construction. Every later theorem will prove one arrow in this route.

2.1 Reduction direction and construction direction

Suppose $M = pm$, where p is prime and $m < M$. The central reduction is

$$\boxed{\text{generator for the modulus-}m\text{ problem} \quad \Longrightarrow \quad \text{generator for the modulus-}pm\text{ problem.}}$$

Thus the proof reduces a larger modulus to a smaller one. Repeatedly removing prime factors gives

$$M \longrightarrow M/p_1 \longrightarrow M/(p_1p_2) \longrightarrow \cdots \longrightarrow 1.$$

The modulus-1 problem is trivial. The generator is assembled in the reverse direction: start from modulus 1 and insert one prime factor at a time. Keeping these two directions separate prevents a common source of confusion:

Key takeaway. The analysis moves from larger moduli to smaller moduli, whereas the construction moves from smaller moduli to larger moduli.

2.2 Why a state has two parameters

To insert a factor p, the reduction works over the alphabet $\mathbb{F}_p$. An intermediate problem is therefore described by a pair (p, m), where p is the current alphabet and m is the current test modulus. We say that the state (p, m) is solved when we have an explicit source on $\mathbb{F}_p^n$ that fools the relevant modulus-m tests with the required error and seed length. The precise tests are defined in Section 3.

The target state is $(2, M)$, because $\mathbb{F}_2^n = \{0,1\}^n$. A starting state $(p, 1)$ costs no randomness, since every modulus-1 phase is the constant function 1.

2.3 The two moves

LIFT: insert one matching prime. If state (p, m) has been solved, LIFT_p constructs state (p, pm):

$$\boxed{\mathrm{LIFT}_p : (p, m) \longrightarrow (p, pm).}$$

The alphabet stays $\mathbb{F}_p$ and the modulus gains one factor p. The construction adds $p-1$ independent small-bias random vectors, one for each finite-difference level. Section 4 proves the error bound.

TRANSFER: move to the next alphabet. If $p < r$ and state (r, m) has been solved, $\mathrm{TRANSFER}_{r\to p}$ constructs state (p, m):

$$\boxed{\mathrm{TRANSFER}_{r\to p} : (r, m) \longrightarrow (p, m).}$$

The modulus stays fixed. The construction applies a coordinatewise map $\mathbb{F}_r \to \mathbb{F}_p$ that reaches every element of $\mathbb{F}_p$ and adds one independent small-bias random vector. Section 5 gives the proof.

2.4 Processing the prime factors

Write the prime factorization with multiplicity as

$$M = p_1 p_2 \cdots p_k, \qquad p_1 \le p_2 \le \cdots \le p_k.$$

The construction processes the factors from largest to smallest. It begins at $(p_k, 1)$ and first applies LIFT_{p_k}. Suppose the factors $p_{j+1}, \ldots, p_k$ have already been inserted, and let $m_j = p_{j+1} \cdots p_k$.

If $p_j = p_{j+1}$, the alphabet already matches the next factor, so apply

$$(p_j, m_j) \xrightarrow{\mathrm{LIFT}_{p_j}} (p_j, p_j m_j).$$

If $p_j < p_{j+1}$, first change alphabets and then insert the factor:

$$(p_{j+1}, m_j) \xrightarrow{\mathrm{TRANSFER}_{p_{j+1}\to p_j}} (p_j, m_j) \xrightarrow{\mathrm{LIFT}_{p_j}} (p_j, p_j m_j).$$

After all factors are inserted, the state is (p_1, M). If $p_1 > 2$, one final transfer reaches $(2, M)$.

2.5 Example: modulus 15

For $15 = 5 \cdot 3$, the complete route is

$$\boxed{(5,1) \xrightarrow{\mathrm{LIFT}_5} (5,5) \xrightarrow{\mathrm{TRANSFER}_{5\to 3}} (3,5) \xrightarrow{\mathrm{LIFT}_3} (3,15) \xrightarrow{\mathrm{TRANSFER}_{3\to 2}} (2,15).}$$

The first and third arrows insert prime factors. The second and fourth arrows change the alphabet without changing the modulus. The route starts from a test class that is identically 1 and ends at the Boolean modulus-15 class.

What this section establishes.
The proof is reduced to two statements, LIFT_p and $\mathrm{TRANSFER}_{r\to p}$. Since M is fixed, the number of arrows is a constant depending only on M. Each arrow will require a constant number of small-bias random vectors and will preserve a seed bound of the form $O_M(\log n + \log(1/\varepsilon))$.

3 The intermediate test classes

Section 2 described the route using states (p, m). We now define exactly which tests belong to a state, what it means for a source to fool them, and the Fourier facts used later. At the end of the section, LIFT and $\mathrm{TRANSFER}$ will be precise mathematical claims rather than informal arrows.

3.1 From modular sums to phase functions

Throughout, n denotes the output length and U_n the uniform distribution on $\{0,1\}^n$. For random variables X, Y taking values in a finite set Ω, the statistical distance is

$$\mathrm{dist}(X,Y) := \frac{1}{2}\sum_{z\in\Omega} \big|\Pr[X=z] - \Pr[Y=z]\big|.$$

Fix an integer $M \ge 2$. For a coefficient vector $a \in \mathbb{Z}_M^n$ consider the modular sum

$$S_a^{(M)}(x) := \sum_{i=1}^{n} a_i x_i \pmod{M}, \qquad x \in \{0,1\}^n.$$

A bit distribution X *ε-fools Boolean linear sums modulo M* if

$$\mathrm{dist}\big(S_a^{(M)}(X),\, S_a^{(M)}(U_n)\big) \le \varepsilon \qquad \text{for every } a \in \mathbb{Z}_M^n.$$

Let

$$\omega_M := e^{2\pi i/M}.$$

The map $r \mapsto \omega_M^r$ translates a residue into a point on the complex unit circle, and is useful because

$$u \equiv v \pmod{M} \iff \omega_M^u = \omega_M^v, \qquad \omega_M^{u+v} = \omega_M^u \omega_M^v.$$

For a fixed coefficient vector a, the associated *phase function* is

$$\boxed{\phi_a^{(M)}(x) := \omega_M^{S_a^{(M)}(x)} = \omega_M^{\sum_i a_i x_i}, \qquad x \in \mathbb{F}_2^n.}$$

Object type: the phase is a function, not a distribution. The map $\phi_a^{(M)} : \mathbb{F}_2^n \to \mathbb{C}$ is a general complex-valued function. Its values have modulus 1; they need not be nonnegative and do not sum to 1. It is therefore not a probability distribution. Passing from residues to phases turns modular arithmetic into multiplicative algebra, and turns balance of the residues into cancellation of a complex expectation.

The one feature of $\phi_a^{(M)}$ that drives everything below is that it *factorises across coordinates*:

$$\phi_a^{(M)}(x) = \prod_{i=1}^{n} \omega_M^{a_i x_i}.$$

Each factor depends on a single input coordinate. The class of tests we shall actually fool is obtained by keeping this product structure and relaxing everything else: the exponent $a_i x_i$ is replaced by an arbitrary function of x_i, and the input alphabet $\{0, 1\}$ is replaced by an arbitrary prime field.

3.2 The test class for an intermediate state

Section 3.1 replaced a modular sum by its phase function

$$\phi_a^{(M)}(x) = \omega_M^{\sum_i a_i x_i} = \prod_{i=1}^{n} \omega_M^{a_i x_i}.$$

This representation exposes the one structural feature that the recursive construction will preserve throughout the paper: the test is a product of contributions from the individual coordinates.

However, the special exponent $a_i x_i$ is too rigid for the intermediate stages of the proof. The construction will perform two operations on the input tests.

First operation: change the alphabet. In the TRANSFER step, a coordinate over one prime field is replaced by a function of a coordinate over another prime field. If

$$\tau : \mathbb{F}_r \longrightarrow \mathbb{F}_p$$

is the coordinate map, then an exponent of the form $a_i x_i$ becomes $a_i \tau(y_i)$. This still depends on only one coordinate, but it need not be linear in that coordinate.

Second operation: compare two shifted copies of a phase. In the LIFT step, we will compare two shifted copies of the same phase. On a single coordinate, this replaces an exponent map q_i by an expression of the form

$$q_i(z+u) - q_i(z+v).$$

Again, the result is an arbitrary function of one coordinate, even if the original exponent was much simpler.

These two observations lead to a single natural choice:

$$\boxed{\text{keep the coordinatewise product structure, but allow arbitrary local exponent maps.}}$$

This enlargement is exactly what we need. It still contains the original Boolean modular phases, and it is stable under both operations used by the construction.

Definition 3.1 (Local phase class). *Let p be a prime and $m \geq 1$ an integer. For a tuple $\mathbf{q} = (q_1, \dots, q_n)$ of* local exponent maps

$$q_i : \mathbb{F}_p \longrightarrow \mathbb{Z}_m,$$

define the associated local phase *by*

$$\Phi_{\mathbf{q}}(x) := \omega_m^{\sum_{i=1}^n q_i(x_i)}, \qquad x \in \mathbb{F}_p^n.$$

Equivalently,

$$\Phi_{\mathbf{q}}(x) = \prod_{i=1}^n \omega_m^{q_i(x_i)}.$$

The class of all such phases is denoted

$$\boxed{\mathsf{Local}_n(p, m)}.$$

For $m = 1$ we use the convention $\mathbb{Z}_1 = \{0\}$ and $\omega_1 = 1$, so $\mathsf{Local}_n(p, 1)$ contains only the constant function 1.

The two parameters play different roles. The prime p specifies the *alphabet* of the input $x \in \mathbb{F}_p^n$, while m specifies the *modulus* of the exponent and hence the roots of unity taken by the phase. No relation between p and m is assumed.

We now verify the three reasons this is the right test class.

Remark 3.2 (Modular sums are local phases). *Let $a \in \mathbb{Z}_M^n$. Define $q_i(0) := 0$ and $q_i(1) := a_i$. Since $q_i(x_i) = a_i x_i$ for $x_i \in \mathbb{F}_2 = \{0, 1\}$, we obtain*

$$\Phi_{\mathbf{q}}(x) = \omega_M^{\sum_i a_i x_i} = \phi_a^{(M)}(x).$$

Hence every Boolean modular phase lies in $\mathsf{Local}_n(2, M)$. Fooling $\mathsf{Local}_n(2, M)$ therefore suffices for the main theorem, and this is why the final target state of the paper is $(2, M)$.

Remark 3.3 (Local phases are stable when the alphabet changes)**.** *Let $\tau : \mathbb{F}_r \to \mathbb{F}_p$ be any map and let $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p,m)$. Then*

$$\Phi_{\mathbf{q}}\big(\tau(y_1),\ldots,\tau(y_n)\big) = \omega_m^{\sum_i q_i(\tau(y_i))} = \omega_m^{\sum_i (q_i\circ\tau)(y_i)}.$$

Each composition $q_i \circ \tau : \mathbb{F}_r \to \mathbb{Z}_m$ is again a valid local exponent map, so the resulting function belongs to $\mathsf{Local}_n(r,m)$. This is the closure property used by TRANSFER*: the proof may change alphabets without leaving the family of tests being fooled.*

Remark 3.4 (Local phases are stable under shifted comparison)**.** *Let $u, v \in \mathbb{F}_p^n$ and $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p,m)$. Then*

$$\begin{aligned}\Phi_{\mathbf{q}}(x+u)\overline{\Phi_{\mathbf{q}}(x+v)} &= \omega_m^{\sum_i q_i(x_i+u_i) - \sum_i q_i(x_i+v_i)}\\ &= \omega_m^{\sum_i \big(q_i(x_i+u_i) - q_i(x_i+v_i)\big)}.\end{aligned}$$

Thus the comparison of two shifted copies is again a local phase, with new local exponent maps

$$q_i'(z) := q_i(z+u_i) - q_i(z+v_i).$$

This is the closure property used by LIFT*. In Section 4 we will introduce this operation formally as the* multiplicative derivative *and iterate it to simplify modulus-pm phases.*

Definition 3.5 (Error against a class)**.** *For a distribution X on $\mathbb{F}_p^n$ and a class $\mathcal{C}$ of functions $\mathbb{F}_p^n \to \mathbb{C}$, define*

$$\boxed{\mathrm{Err}_X(\mathcal{C}) := \sup_{f\in\mathcal{C}} \Big|\mathbb{E}\left[f(X)\right] - \mathbb{E}\left[f(U_p^n)\right]\Big|.}$$

*We say that X ε-*fools *$\mathcal{C}$ when $\mathrm{Err}_X(\mathcal{C}) \le \varepsilon$.*

This quantity gives a single numerical meaning to the phrase "the source fools the state." It will be used uniformly throughout the recursion: it measures how well the best test in the class distinguishes X from the uniform input. As usual, if $\mathcal{C}' \subseteq \mathcal{C}$, then $\mathrm{Err}_X(\mathcal{C}') \le \mathrm{Err}_X(\mathcal{C})$.

3.3 The precise meaning of a solved state

State (p,m) is solved with error ε if there is an explicit source X on $\mathbb{F}_p^n$ such that

$$\mathrm{Err}_X\big(\mathsf{Local}_n(p,m)\big) \le \varepsilon$$

and whose seed length is

$$O_{p,m}\left(\log n + \log\frac{1}{\varepsilon}\right).$$

The starting state $(p,1)$ is solved by the constant-zero source with error zero. The target state is $(2,M)$; by Remark 3.2, it contains every Boolean modular phase needed for the theorem.

With this definition, the remaining construction claims are exactly

$$\begin{aligned}\mathrm{LIFT}_p &: \text{solve } (p,m) \text{ sufficiently accurately and obtain } (p,pm),\\ \mathrm{TRANSFER}_{r\to p} &: \text{solve } (r,m) \text{ sufficiently accurately and obtain } (p,m).\end{aligned}$$

The rest of the paper proves these statements and tracks the required accuracy.

3.4 Fourier analysis and small-bias random vectors

Both moves are analysed by Fourier analysis over the input group $\mathbb{F}_p^n$, and the final conversion to statistical distance uses Fourier analysis over $\mathbb{Z}_M$. We record the common framework once and then specialise.

Let H be a finite abelian group, written additively. A *character* of H is a function $\chi : H \to \mathbb{C}$ with $\chi(x + y) = \chi(x)\chi(y)$ for all $x, y \in H$; a character converts the group operation into multiplication. Since H is finite, $|\chi(x)| = 1$ and $\chi(-x) = \overline{\chi(x)}$. The characters form a group $\widehat{H}$ under pointwise multiplication, with $|\widehat{H}| = |H|$, and they are orthonormal for the normalised inner product $\langle f, g\rangle := \mathbb{E}_{x\sim U_H}[f(x)\overline{g(x)}]$. Being $|H|$ in number, they form an orthonormal basis of all functions $H \to \mathbb{C}$.

Conceptual meaning of the Fourier transform. The characters are multiplicative with respect to the group operation, while the Fourier expansion is linear. Thus Fourier analysis represents a function as a linear combination of multiplicative basis functions.

We now fix the input group. For a prime p write $\omega_p = e^{2\pi i/p}$, and for $\xi, x \in \mathbb{F}_p^n$ let

$$\chi_\xi(x) := \omega_p^{\langle \xi, x\rangle}, \qquad \langle \xi, x\rangle = \sum_{i=1}^{n} \xi_i x_i \pmod p.$$

These are all the characters of $\mathbb{F}_p^n$. For $f : \mathbb{F}_p^n \to \mathbb{C}$ define

$$\widehat{f}(\xi) := \mathbb{E}_{x\sim U_p^n}\left[f(x)\overline{\chi_\xi(x)}\right], \qquad \text{so that} \qquad f(x) = \sum_{\xi\in\mathbb{F}_p^n} \widehat{f}(\xi)\chi_\xi(x),$$

and Parseval's identity holds:

$$\boxed{\sum_{\xi\in\mathbb{F}_p^n} |\widehat{f}(\xi)|^2 = \mathbb{E}_{x\sim U_p^n}|f(x)|^2.}$$

In particular, if $|f| = 1$ everywhere then $\sum_\xi |\widehat{f}(\xi)|^2 = 1$, a normalisation used at several points below. Note also that $\widehat{f}(0) = \mathbb{E}f(U_p^n)$: the trivial Fourier coefficient of a test is its uniform expectation, which is precisely the quantity that Definition 3.5 compares against.

Definition 3.6 (Small bias). *A distribution B on $\mathbb{F}_p^n$ is* δ-biased *if*

$$|\mathbb{E}\chi_\xi(B)| \le \delta \qquad \textit{for every } \xi \neq 0.$$

Imported primitive: small bias over a fixed prime field. For every fixed prime p, every n, and every $0 < \delta \le 1$ there is an explicit δ-biased distribution on $\mathbb{F}_p^n$ with seed length

$$O_p\left(\log n + \log\frac{1}{\delta}\right)$$

[NN93], [AGHP92], [AMN98]. This is the only pseudorandomness primitive the upper bound imports; everything else is proved here.

Two consequences of the definitions will be used in both moves. First, adding an independent random vector multiplies each character expectation: if B is independent of X then, by multiplicativity of characters,

$$\mathbb{E}\chi_\xi(X+B) = \mathbb{E}\chi_\xi(X) \cdot \mathbb{E}\chi_\xi(B),$$

so in particular $X + B$ is δ-biased whenever B is, irrespective of X. Second, and equivalently, if $h(x) := \mathbb{E}_B[f(x+B)]$ denotes the smoothing of f by B, then

$$\boxed{\widehat{h}(\xi) = \widehat{f}(\xi) \cdot \mathbb{E}\chi_\xi(B).}$$

Smoothing by a small-bias random vector thus attenuates every nontrivial Fourier coefficient by a factor of at most δ, and leaves the trivial coefficient untouched.

Summary of Section 3. The definitions introduced in this section fit together as follows:

$\phi_a^{(M)}$ $\leftarrow$ encode a modular sum as a multiplicative phase, so the test factorises by coordinates;

$\mathsf{Local}_n(p,m)$ $\leftarrow$ enlarge the test class while preserving that coordinatewise structure, so it survives both LIFT and TRANSFER;

$\mathrm{Err}_X(\mathsf{Local}_n(p,m))$ $\leftarrow$ measure how well a source fools that class, giving a numerical meaning to a solved state;

(p,m) $\leftarrow$ the intermediate goal "fool $\mathsf{Local}_n(p,m)$ over $\mathbb{F}_p^n$," which the recursion moves between;

$\widehat{f}$, δ-bias $\leftarrow$ the Fourier language and smoothing primitive, which power the later LIFT and TRANSFER proofs.

In short: Section 3.1 defines the tests, Section 3.3 defines what it means to fool them at an intermediate state, and Section 3.4 records the Fourier tools that will drive the two construction moves.

4 The LIFT reduction

The state (p,m) is assumed to be solved, and the next goal is to solve the state (p,pm). The alphabet remains $\mathbb{F}_p$; only the modulus gains a factor p.

The proof compares a phase at two shifted inputs. We first define this operation and explain how it acts on local phases. We then state the two-step plan for constructing the new generator.

4.1 Phase derivatives

Let $f : \mathbb{F}_p^n \to \mathbb{C}$ have absolute value 1 everywhere. For shifts $u, v \in \mathbb{F}_p^n$, define the *multiplicative derivative* of f by

$$D_{u,v}f(x) := f(x+u)\overline{f(x+v)}. \tag{4.1}$$

If

$$f(x) = \omega_m^{P(x)},$$

then

$$\begin{aligned} D_{u,v} f(x) &= \omega_m^{P(x+u)} \overline{\omega_m^{P(x+v)}} \\ &= \omega_m^{P(x+u)-P(x+v)}. \end{aligned}$$

Thus the multiplicative derivative of a phase corresponds to the ordinary finite difference of its exponent. We use the ratio of the two shifted phases rather than their ordinary difference because the ratio is again a phase.

For a class $\mathcal{C}$ of unit-modulus functions, define

$$\partial \mathcal{C} := \left\{ D_{u,v} f : f \in \mathcal{C},\ u, v \in \mathbb{F}_p^n \right\}.$$

We also define the repeated derivative classes recursively by

$$\partial^0 \mathcal{C} := \mathcal{C}, \qquad \partial^{j+1} \mathcal{C} := \partial\big(\partial^j \mathcal{C}\big).$$

Multiplicative derivatives preserve the coordinatewise form of a local phase. Indeed, let

$$\Phi_{\mathbf{q}}(x) = \omega_m^{\sum_{i=1}^n q_i(x_i)} \in \mathsf{Local}_n(p, m).$$

Then

$$\begin{aligned} D_{u,v} \Phi_{\mathbf{q}}(x) &= \omega_m^{\sum_i q_i(x_i+u_i) - \sum_i q_i(x_i+v_i)} \\ &= \omega_m^{\sum_i (q_i(x_i+u_i) - q_i(x_i+v_i))}. \end{aligned}$$

Hence

$$D_{u,v} \Phi_{\mathbf{q}} = \Phi_{\mathbf{q}'},$$

where the new local exponent maps are

$$q_i'(z) := q_i(z+u_i) - q_i(z+v_i).$$

In particular,

$$\partial \mathsf{Local}_n(p, m) \subseteq \mathsf{Local}_n(p, m),$$

and the same remains true after any number of repeated derivatives.

4.2 The two-step plan for LIFT_p

We can now state the two parts of the proof precisely.

Proof strategy. Let X_0 be a source that fools $\mathsf{Local}_n(p, m)$.

Step 1: Descend from modulus pm to modulus m. We prove that

$$\mathrm{Err}_{X_0}\big(\mathsf{Local}_n(p, m)\big) \le \alpha \quad \Longrightarrow \quad \mathrm{Err}_{X_0}\big(\partial^{p-1} \mathsf{Local}_n(p, pm)\big) \le \alpha.$$

The reason is that every phase in

$$\partial^{p-1} \mathsf{Local}_n(p, pm)$$

is a constant complex factor times a phase in $\mathsf{Local}_n(p, m)$.

To prove this, we reduce each local exponent modulo p and represent the resulting function on $\mathbb{F}_p$ by a polynomial of degree at most $p-1$. Each nonzero finite difference lowers the degree by one, so after $p-1$ derivatives the exponent is constant modulo p. Its remaining input-dependent part is therefore divisible by p, which changes the modulus from pm to m.

Step 2: Recover the original target class. We prove a one-step lifting statement: if a source X fools $\partial\mathcal{C}$, then adding one independent small-bias random vector B produces a source $X+B$ that fools $\mathcal{C}$, with a controlled loss in the error.

Starting from the source X_0, which after Step 1 fools

$$\partial^{p-1}\mathsf{Local}_n(p, pm),$$

we apply this one-step statement $p-1$ times. At each step we add one independent small-bias random vector and move upward by one derivative level:

$$\partial^{p-1}\mathsf{Local}_n(p, pm) \longrightarrow \partial^{p-2}\mathsf{Local}_n(p, pm) \longrightarrow \cdots \longrightarrow \partial\mathsf{Local}_n(p, pm) \longrightarrow \mathsf{Local}_n(p, pm).$$

The final source therefore fools $\mathsf{Local}_n(p, pm)$ and solves the state (p, pm).

4.3 Why $p-1$ derivatives remove one factor p

The algebraic point is easiest to see after reducing an exponent modulo p. We use the following standard fact.

Lemma 4.1 (Polynomial representation of functions on $\mathbb{F}_p$)**.** *Every function $g : \mathbb{F}_p \to \mathbb{F}_p$ is represented on $\mathbb{F}_p$ by a polynomial $G \in \mathbb{F}_p[z]$ of degree at most $p-1$.*

Proof. For each $a \in \mathbb{F}_p$, let

$$L_a(z) = \prod_{b \in \mathbb{F}_p \setminus \{a\}} \frac{z-b}{a-b}.$$

Then $L_a(a) = 1$ and $L_a(b) = 0$ for $b \neq a$. Hence

$$G(z) = \sum_{a \in \mathbb{F}_p} g(a) L_a(z)$$

has degree at most $p-1$ and satisfies $G(a) = g(a)$ for every $a \in \mathbb{F}_p$. □

Lemma 4.2 (A nonzero finite difference lowers the degree)**.** *Let $G \in \mathbb{F}_p[z]$ have degree d with $1 \leq d \leq p-1$, and let $u, v \in \mathbb{F}_p$ with $u \neq v$. Then*

$$G(z+u) - G(z+v)$$

has degree $d-1$. If $u = v$, the difference is the zero polynomial.

Proof. Write the leading term of G as $a_d z^d$. The coefficient of z^{d-1} in $G(z+u) - G(z+v)$ is

$$a_d d(u-v).$$

It is nonzero in $\mathbb{F}_p$ because $a_d \neq 0$, $1 \leq d \leq p-1$, and $u \neq v$. All terms of degree d cancel, so the resulting degree is exactly $d-1$. The case $u = v$ is immediate. □

Proposition 4.3 (Prime-factor descent)**.** *Let p be prime and $m \geq 1$. Every function in $\partial^{p-1}\mathsf{Local}_n(p, pm)$ has the form*

$$\lambda\Psi, \qquad |\lambda| = 1, \qquad \Psi \in \mathsf{Local}_n(p, m).$$

Consequently, any source that α-fools $\mathsf{Local}_n(p, m)$ also α-fools $\partial^{p-1}\mathsf{Local}_n(p, pm)$.

Proof. Take

$$\Phi_{\mathbf{q}}(x) = \omega_{pm}^{\sum_{i=1}^{n} q_i(x_i)} \in \mathsf{Local}_n(p, pm),$$

where

$$\mathbf{q} = (q_1, \dots, q_n), \qquad q_i : \mathbb{F}_p \to \mathbb{Z}_{pm}.$$

Apply any sequence of $p - 1$ multiplicative derivatives to $\Phi_{\mathbf{q}}$. Since derivatives preserve the local-phase form, the resulting function can be written as

$$F(x) = \omega_{pm}^{\sum_{i=1}^{n} q_i^*(x_i)}$$

for some final exponent maps

$$q_i^* : \mathbb{F}_p \to \mathbb{Z}_{pm}.$$

We show that each q_i^* has the form

$$q_i^*(z) = c_i + p\, r_i(z), \qquad r_i : \mathbb{F}_p \to \mathbb{Z}_m.$$

Step 1: Track one coordinate modulo p. Fix a coordinate i. Reduce the original exponent map modulo p:

$$\overline{q}_i(z) := q_i(z) \pmod{p}, \qquad \overline{q}_i : \mathbb{F}_p \to \mathbb{F}_p.$$

By Lemma 4.1, there is a polynomial

$$P_i(z) \in \mathbb{F}_p[z], \qquad \deg P_i \leq p - 1,$$

such that

$$P_i(z) = \overline{q}_i(z) \qquad \text{for every } z \in \mathbb{F}_p.$$

Suppose one of the derivatives uses shifts $u, v \in \mathbb{F}_p$ in coordinate i. Modulo p, the exponent map is then replaced by

$$P_i(z + u) - P_i(z + v).$$

If $u = v$, this is the zero polynomial. If $u \neq v$, then Lemma 4.2 shows that every nonconstant polynomial loses one degree.

Since the initial degree is at most $p - 1$, after $p - 1$ derivatives the resulting polynomial is constant. Therefore the final exponent map q_i^* is constant modulo p: there exists $c_i \in \mathbb{Z}_{pm}$ such that

$$q_i^*(z) \equiv c_i \pmod{p} \qquad \text{for every } z \in \mathbb{F}_p.$$

Step 2: Factor p out of the remaining input-dependent part. The previous equation means that

$$q_i^*(z) - c_i$$

is a multiple of p in $\mathbb{Z}_{pm}$ for every z. The multiples of p in $\mathbb{Z}_{pm}$ are exactly

$$\{\, pr : r \in \mathbb{Z}_m \,\}.$$

Hence, for every $z \in \mathbb{F}_p$, there is a unique $r_i(z) \in \mathbb{Z}_m$ such that

$$q_i^*(z) - c_i = p\, r_i(z) \qquad \text{in } \mathbb{Z}_{pm}.$$

Thus

$$\boxed{q_i^*(z) = c_i + p\, r_i(z)} \qquad \text{for some } r_i : \mathbb{F}_p \to \mathbb{Z}_m.$$

Step 3: Reassemble the phase. Applying the previous argument to every coordinate gives

$$\begin{aligned} F(x) &= \omega_{pm}^{\sum_i q_i^*(x_i)} \\ &= \omega_{pm}^{\sum_i (c_i + p\, r_i(x_i))} \\ &= \omega_{pm}^{\sum_i c_i}\, \omega_{pm}^{p \sum_i r_i(x_i)} \\ &= \omega_{pm}^{\sum_i c_i}\, \omega_{m}^{\sum_i r_i(x_i)}. \end{aligned}$$

Where the last inequality is true because for every p, m we have $\omega_{pm}^p = \omega_m$. Define

$$\lambda := \omega_{pm}^{\sum_i c_i}, \qquad \Psi(x) := \omega_m^{\sum_i r_i(x_i)}.$$

Then

$$|\lambda| = 1 \qquad \text{and} \qquad \Psi \in \mathsf{Local}_n(p, m),$$

so

$$F = \lambda \Psi.$$

Since the original phase and the $p-1$ derivatives were arbitrary, every function in $\partial^{p-1}\mathsf{Local}_n(p, pm)$ has this form.

Finally, suppose that X α-fools $\mathsf{Local}_n(p, m)$. For every $F = \lambda\Psi \in \partial^{p-1}\mathsf{Local}_n(p, pm)$,

$$\begin{aligned} \left|\mathbb{E}F(X) - \mathbb{E}F(U_p^n)\right| &= \left|\lambda \left(\mathbb{E}\Psi(X) - \mathbb{E}\Psi(U_p^n)\right)\right| \\ &= \left|\mathbb{E}\Psi(X) - \mathbb{E}\Psi(U_p^n)\right| \\ &\le \alpha, \end{aligned}$$

because $|\lambda| = 1$ and $\Psi \in \mathsf{Local}_n(p, m)$. Therefore X also α-fools $\partial^{p-1}\mathsf{Local}_n(p, pm)$. $\square$

4.4 One small-bias vector lifts one derivative level

The previous subsection moves downward among test classes. We now prove the opposite movement for the source: if a source fools the derivative class, then adding one small-bias random vector makes it fool the original class.

Lemma 4.4 (Derivative lifting). *Let $\mathcal{C}$ be a class of unit-modulus functions on $\mathbb{F}_p^n$. Suppose X is a distribution on $\mathbb{F}_p^n$ satisfying*

$$\mathrm{Err}_X(\partial\mathcal{C}) \le \alpha,$$

and let B be an independent δ-biased distribution on $\mathbb{F}_p^n$. Then

$$\boxed{\mathrm{Err}_{X+B}(\mathcal{C}) \le \sqrt{3\alpha + \delta^2}.}$$

Proof. Let

$$Y := X + B$$

be the new source. Fix a test $f \in \mathcal{C}$, let

$$U \sim U_p^n, \qquad \mu := \mathbb{E}_U\left[f(U)\right],$$

and denote the error of Y on this particular test by

$$e_f := \mathbb{E}\left[f(Y)\right] - \mu.$$

We prove

$$|e_f|^2 \le 3\alpha + \delta^2.$$

Define the smoothed test

$$\widetilde{f}(x) := \mathbb{E}_B\left[f(x + B)\right].$$

Since

$$\mathbb{E}_X\left[\widetilde{f}(X)\right] = \mathbb{E}\left[f(Y)\right],$$

the error e_f can be written as

$$e_f = \mathbb{E}_X\left[\widetilde{f}(X)\right] - \mu.$$

By Cauchy–Schwarz,

$$\begin{aligned} |e_f|^2 &= \left|\mathbb{E}_X\left[\widetilde{f}(X) - \mu\right]\right|^2 \\ &\le \mathbb{E}_X\left[|\widetilde{f}(X) - \mu|^2\right]. \end{aligned}$$

Expanding the square and using $\mathbb{E}_X\left[\widetilde{f}(X)\right] = \mu + e_f$,

$$\mathbb{E}_X\left[|\widetilde{f}(X) - \mu|^2\right] = \mathbb{E}_X\left[|\widetilde{f}(X)|^2\right] - |\mu|^2 - 2\,\mathrm{Re}(\overline{\mu} e_f).$$

Thus, to bound $|e_f|^2$, it remains to bound the two quantities

$$\mathbb{E}_X\left[|\widetilde{f}(X)|^2\right] \qquad \text{and} \qquad |\overline{\mu} e_f|.$$

We establish these two bounds below.

First, Y still fools the derivative class. Let $g \in \partial\mathcal{C}$. For every fixed $b \in \mathbb{F}_p^n$, the shifted function

$$x \longmapsto g(x + b)$$

also belongs to $\partial\mathcal{C}$. Indeed, if $g = D_{u,v} f_0$ for some $f_0 \in \mathcal{C}$, then

$$g(x+b) = D_{u+b,v+b} f_0(x).$$

Therefore the hypothesis on X gives

$$|\mathbb{E}_X [g(X+b)] - \mathbb{E}_U [g(U)]| \leq \alpha,$$

where we used that $U + b$ is again uniform. Averaging over $b = B$ gives

$$|\mathbb{E}_Y [g(Y)] - \mathbb{E}_U [g(U)]| \leq \alpha \qquad \text{for every } g \in \partial\mathcal{C}. \tag{4.2}$$

We next obtain a correlation bound. Let U' be an independent uniform vector in $\mathbb{F}_p^n$. Fix any $y \in \mathbb{F}_p^n$. Since $y + U'$ is uniform on $\mathbb{F}_p^n$,

$$\mathbb{E}_{U'} \left[f(y)\overline{f(y+U')} \right] = f(y)\overline{\mathbb{E}\left[f(U')\right]} = f(y)\overline{\mu}.$$

Now average this identity over the random vector Y. We obtain

$$\mathbb{E}_{Y,U'} \left[f(Y)\overline{f(Y+U')} \right] = \mathbb{E}\left[f(Y)\right] \overline{\mu}.$$

Similarly, averaging the same identity over $y = U$, where $U \sim U_p^n$, gives

$$\begin{aligned} \mathbb{E}_{U,U'} \left[f(U)\overline{f(U+U')} \right] &= \mathbb{E}\left[f(U)\right] \overline{\mu} \\ &= \mu\overline{\mu}. \end{aligned}$$

Subtracting the two identities gives

$$e_f\overline{\mu} = \mathbb{E}_{Y,U'} \left[f(Y)\overline{f(Y+U')} \right] - \mathbb{E}_{U,U'} \left[f(U)\overline{f(U+U')} \right].$$

Recall that for every fixed shift $u \in \mathbb{F}_p^n$,

$$D_{0,u} f(x) = f(x)\overline{f(x+u)}.$$

Therefore, if we first fix u and then average over the input, the two terms above become

$$\mathbb{E}_Y \left[D_{0,u} f(Y)\right] \qquad \text{and} \qquad \mathbb{E}_U \left[D_{0,u} f(U)\right].$$

Define the error for this derivative by

$$c_f(u) := \mathbb{E}_Y \left[D_{0,u} f(Y)\right] - \mathbb{E}_U \left[D_{0,u} f(U)\right].$$

Since $D_{0,u} f \in \partial\mathcal{C}$, equation (4.2) gives

$$|c_f(u)| \leq \alpha \qquad \text{for every } u \in \mathbb{F}_p^n.$$

Averaging over the uniform shift U' now gives

$$e_f\overline{\mu} = \mathbb{E}_{U'}\left[c_f(U')\right].$$

Hence, by the triangle inequality,

$$\boxed{|e_f\overline{\mu}| \le \mathbb{E}_{U'}\left[|c_f(U')|\right] \le \alpha.}$$

$$\boxed{|e_f\overline{\mu}| \le \alpha.} \tag{4.3}$$

We now bound the second moment of the smoothed test. Let B' be an independent copy of B. Then

$$\begin{aligned}|\widetilde{f}(x)|^2 &= \mathbb{E}_{B,B'}\left[f(x+B)\overline{f(x+B')}\right]\\ &= \mathbb{E}_{B,B'}\left[D_{B,B'}f(x)\right].\end{aligned}$$

For every fixed b, b', the function $D_{b,b'}f$ belongs to $\partial\mathcal{C}$. Therefore

$$\mathbb{E}_X\left[|\widetilde{f}(X)|^2\right] \le \mathbb{E}_U\left[|\widetilde{f}(U)|^2\right] + \alpha.$$

It remains to bound the second term. By Fourier expansion,

$$\widehat{\widetilde{f}}(\xi) = \widehat{f}(\xi)\mathbb{E}\left[\chi_\xi(B)\right].$$

Hence Parseval and the δ-bias of B give

$$\begin{aligned}\mathbb{E}_U\left[|\widetilde{f}(U)|^2\right] &= \sum_{\xi\in\mathbb{F}_p^n} |\widehat{f}(\xi)|^2\,|\mathbb{E}\left[\chi_\xi(B)\right]|^2\\ &\le |\widehat{f}(0)|^2 + \delta^2\sum_{\xi\neq 0}|\widehat{f}(\xi)|^2\\ &\le |\mu|^2 + \delta^2,\end{aligned}$$

where $\widehat{f}(0) = \mu$ and

$$\sum_{\xi}|\widehat{f}(\xi)|^2 = 1$$

because $|f| = 1$. Therefore

$$\boxed{\mathbb{E}_X\left[|\widetilde{f}(X)|^2\right] \le |\mu|^2 + \alpha + \delta^2.} \tag{4.4}$$

Finally, combine the two bounds. Using (4.4) and (4.3),

$$\begin{aligned}|e_f|^2 &\le \alpha + \delta^2 - 2\,\mathrm{Re}(\overline{\mu}e_f)\\ &\le \alpha + \delta^2 + 2|\overline{\mu}e_f|\\ &\le 3\alpha + \delta^2.\end{aligned}$$

Thus

$$|e_f| \le \sqrt{3\alpha + \delta^2}.$$

Since $f \in \mathcal{C}$ was arbitrary, the lemma follows. $\square$

4.5 The LIFT theorem

Theorem 4.5 (LIFT)**.** *Let p be prime, $m \geq 1$, and $0 < \varepsilon \leq 1$. Set*

$$\theta := \left(\frac{\varepsilon}{4}\right)^{2^{p-1}}.$$

Let X_0 be a distribution on $\mathbb{F}_p^n$ satisfying

$$\mathrm{Err}_{X_0}(\mathsf{Local}_n(p, m)) \leq \theta,$$

and let $B_1, \ldots, B_{p-1}$ be independent θ-biased distributions on $\mathbb{F}_p^n$, also independent of X_0. Define

$$X_j = X_{j-1} + B_j, \qquad 1 \leq j \leq p-1.$$

Then

$$\boxed{\mathrm{Err}_{X_{p-1}}(\mathsf{Local}_n(p, pm)) \leq \varepsilon.}$$

Moreover, if X_0 has seed length $O_{p,m}(\log n + \log(1/\theta))$, then X_{p-1} has seed length

$$O_{p,m}\left(\log n + \log \frac{1}{\varepsilon}\right).$$

Proof. Define the derivative hierarchy

$$\mathcal{C}_j := \partial^{p-1-j}\mathsf{Local}_n(p, pm), \qquad 0 \leq j \leq p-1.$$

Thus $\mathcal{C}_{p-1} = \mathsf{Local}_n(p, pm)$ and $\partial\mathcal{C}_j = \mathcal{C}_{j-1}$ for $j \geq 1$. By Proposition 4.3, the source X_0 fools $\mathcal{C}_0$ with error at most θ.

Let

$$e_j = \mathrm{Err}_{X_j}(\mathcal{C}_j).$$

Lemma 4.4 gives

$$e_j \leq \sqrt{3e_{j-1} + \theta^2} \qquad (1 \leq j \leq p-1). \tag{4.5}$$

To solve this recursion, define

$$E_0 = \theta, \qquad E_j = \sqrt{3E_{j-1} + \theta^2}.$$

Inductively, $e_j \leq E_j$. Also $E_j \geq \theta$ for every j, and $0 < \theta \leq 1$, so $\theta^2 \leq E_{j-1}$. Therefore

$$E_j \leq \sqrt{4E_{j-1}} = 2\sqrt{E_{j-1}}.$$

A second induction now gives

$$E_j \leq 4\theta^{1/2^j}.$$

For $j = 0$ this is immediate. If it holds for $j-1$, then

$$E_j \leq 2\sqrt{E_{j-1}} \leq 2\sqrt{4\theta^{1/2^{j-1}}} = 4\theta^{1/2^j}.$$

At the final level,

$$e_{p-1} \leq 4\theta^{1/2^{p-1}} = 4 \cdot \frac{\varepsilon}{4} = \varepsilon.$$

For the seed length,

$$\log \frac{1}{\theta} = 2^{p-1} \log \frac{4}{\varepsilon} = O_p\left(1 + \log \frac{1}{\varepsilon}\right).$$

The construction adds $p-1$ independent small-bias vectors, a number depending only on the fixed prime p. Their seed lengths and the seed length of X_0 therefore sum to $O_{p,m}(\log n + \log(1/\varepsilon))$. $\square$

What this section establishes.
The operation LIFT_p is proved:

$$\boxed{(p, m) \xrightarrow{\mathrm{LIFT}_p} (p, pm).}$$

It uses $p-1$ small-bias random vectors. The algebra lowers the test class by $p-1$ derivative levels, and Lemma 4.4 climbs those levels back one at a time.

5 The TRANSFER reduction

The operation LIFT_p can insert a prime factor only when the source alphabet is $\mathbb{F}_p$. After one prime factor has been processed, the construction may need to move to a smaller prime alphabet before inserting the next one. This section proves that changing the alphabet preserves the modulus and the optimal seed-length form.

The state (r, m) is assumed to be solved, and the next goal is to solve the state (p, m), where $p < r$. The modulus remains m.

Let Y be a source on $\mathbb{F}_r^n$ that fools $\mathsf{Local}_n(r, m)$. Fix a map

$$\sigma : \mathbb{F}_r \to \mathbb{F}_p$$

that *reaches every element of* $\mathbb{F}_p$; that is, for every $a \in \mathbb{F}_p$ there exists some $y \in \mathbb{F}_r$ such that $\sigma(y) = a$. Equivalently, if U_r is uniform on $\mathbb{F}_r$, then $\sigma(U_r)$ has full support on $\mathbb{F}_p$, and in fact

$$\Pr[\sigma(U_r) = a] \geq \frac{1}{r} \qquad \text{for every } a \in \mathbb{F}_p.$$

Such a map can be chosen explicitly as follows. Identify

$$\mathbb{F}_r = \{0, 1, \ldots, r-1\}, \qquad \mathbb{F}_p = \{0, 1, \ldots, p-1\},$$

using the standard integer representatives, and define

$$\boxed{\sigma(x) := x \pmod{p}.}$$

Thus σ simply takes the representative of $x \in \mathbb{F}_r$ and reduces it modulo p. Since $p < r$, every $a \in \mathbb{F}_p$ has at least one preimage, namely $x = a$. Therefore σ reaches every element of $\mathbb{F}_p$. More precisely, each $a \in \mathbb{F}_p$ has either

$$\left\lfloor \frac{r}{p} \right\rfloor \qquad \text{or} \qquad \left\lceil \frac{r}{p} \right\rceil$$

preimages. Hence, if U_r is uniform on $\mathbb{F}_r$,

$$\Pr\left[\sigma(U_r) = a\right] = \frac{|\sigma^{-1}(a)|}{r} \geq \frac{1}{r} \qquad \text{for every } a \in \mathbb{F}_p.$$

Thus $\sigma(U_r)$ has full support on $\mathbb{F}_p$, although it is generally not uniform. We apply σ separately to every coordinate of Y, add an independent small-bias vector $B \in \mathbb{F}_p^n$, and define

$$X := \sigma(Y) + B.$$

The difficulty is precisely that even if the input to σ is uniform on $\mathbb{F}_r$, its image under σ need not be uniform on $\mathbb{F}_p$. The proof handles this discrepancy using two different estimates.

Proof strategy. Fix a target test $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, m)$, let $U_p \sim U_p^n$, and let w be the number of active coordinates of $\Phi_{\mathbf{q}}$. We prove two bounds for

$$\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right|.$$

Low-weight estimate. The small-bias vector B makes $X = \sigma(Y) + B$ small-biased, regardless of the distribution of $\sigma(Y)$. A test with w active coordinates has at most p^w nonzero Fourier coefficients, which gives

$$\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \leq \delta p^{w/2}.$$

This estimate is strongest when w is small.

High-weight estimate. Condition on a fixed value $B = \beta$. The function

$$y \longmapsto \Phi_{\mathbf{q}}(\sigma(y) + \beta)$$

belongs to $\mathsf{Local}_n(r, m)$, so the pseudorandomness guarantee for Y applies. After replacing Y by a uniform vector over $\mathbb{F}_r^n$, independence across coordinates implies that every active coordinate contributes a factor of absolute value at most some constant $\kappa < 1$. The same cancellation occurs under uniform input over $\mathbb{F}_p^n$. If the incoming source has error η, this gives

$$\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \leq \eta + 2\kappa^w.$$

This estimate is strongest when w is large.

The first bound increases with w, while the second decreases with w. We choose a threshold W so that the low-weight estimate handles $w \leq W$ and the high-weight estimate handles $w > W$.

5.1 Active coordinates and the low-weight estimate

A coordinate i is *active* for $\mathbf{q}$ when $q_i : \mathbb{F}_p \to \mathbb{Z}_m$ is nonconstant. Let

$$w = \mathrm{wt}(\mathbf{q})$$

be the number of active coordinates. Inactive coordinates contribute only a constant root of unity, so w is the number of coordinates on which the test genuinely depends.

Lemma 5.1 (Low-weight estimate)**.** *Let Z be any distribution on $\mathbb{F}_p^n$, let B be an independent δ-biased distribution, and set $X = Z + B$. For every $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, m)$ of weight w,*

$$\boxed{\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U \sim U_p^n}\left[\Phi_{\mathbf{q}}(U)\right]\right| \le \delta p^{w/2}.}$$

Proof. Let

$$U \sim U_p^n.$$

We first show that the new source

$$X = Z + B$$

is itself δ-biased. Recall that for

$$\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{F}_p^n,$$

the corresponding character is

$$\chi_\xi(x) = \omega_p^{\langle \xi, x\rangle} = \omega_p^{\sum_{i=1}^n \xi_i x_i}.$$

Since characters turn addition into multiplication,

$$\chi_\xi(Z + B) = \chi_\xi(Z)\chi_\xi(B).$$

Because Z and B are independent,

$$\begin{aligned}
\mathbb{E}_X\left[\chi_\xi(X)\right] &= \mathbb{E}_{Z,B}\left[\chi_\xi(Z + B)\right] \\
&= \mathbb{E}_{Z,B}\left[\chi_\xi(Z)\chi_\xi(B)\right] \\
&= \mathbb{E}_Z\left[\chi_\xi(Z)\right]\mathbb{E}_B\left[\chi_\xi(B)\right].
\end{aligned}$$

For every nonzero ξ, the δ-bias of B gives

$$\left|\mathbb{E}_B\left[\chi_\xi(B)\right]\right| \le \delta,$$

while

$$\left|\mathbb{E}_Z\left[\chi_\xi(Z)\right]\right| \le 1$$

because $|\chi_\xi| = 1$. Therefore

$$\boxed{\left|\mathbb{E}_X\left[\chi_\xi(X)\right]\right| \le \delta \qquad \text{for every } \xi \neq 0.}$$

We now apply this fact to the local phase $\Phi_{\mathbf{q}}$. A Fourier coefficient of $\Phi_{\mathbf{q}}$ is indexed by a vector

$$\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{F}_p^n$$

and is defined by

$$\widehat{\Phi_{\mathbf{q}}}(\xi) = \mathbb{E}_U \left[\Phi_{\mathbf{q}}(U) \overline{\chi_\xi(U)} \right].$$

The Fourier expansion is

$$\Phi_{\mathbf{q}}(x) = \sum_{\xi \in \mathbb{F}_p^n} \widehat{\Phi_{\mathbf{q}}}(\xi) \chi_\xi(x).$$

Taking expectations under X gives

$$\mathbb{E}_X \left[\Phi_{\mathbf{q}}(X) \right] = \sum_{\xi \in \mathbb{F}_p^n} \widehat{\Phi_{\mathbf{q}}}(\xi) \mathbb{E}_X \left[\chi_\xi(X) \right].$$

Similarly,

$$\mathbb{E}_U \left[\Phi_{\mathbf{q}}(U) \right] = \sum_{\xi \in \mathbb{F}_p^n} \widehat{\Phi_{\mathbf{q}}}(\xi) \mathbb{E}_U \left[\chi_\xi(U) \right].$$

For uniform U, every nontrivial character has expectation zero:

$$\mathbb{E}_U \left[\chi_\xi(U) \right] = 0 \qquad \text{for every } \xi \neq 0,$$

while for $\xi = 0$,

$$\chi_0(x) = 1$$

under both X and U. Hence the $\xi = 0$ terms cancel exactly, and

$$\mathbb{E}_X \left[\Phi_{\mathbf{q}}(X) \right] - \mathbb{E}_U \left[\Phi_{\mathbf{q}}(U) \right] = \sum_{\xi \neq 0} \widehat{\Phi_{\mathbf{q}}}(\xi) \mathbb{E}_X \left[\chi_\xi(X) \right].$$

Taking absolute values and using the bound above,

$$\begin{aligned} \left| \mathbb{E}_X \left[\Phi_{\mathbf{q}}(X) \right] - \mathbb{E}_U \left[\Phi_{\mathbf{q}}(U) \right] \right| &\leq \sum_{\xi \neq 0} \left| \widehat{\Phi_{\mathbf{q}}}(\xi) \right| \left| \mathbb{E}_X \left[\chi_\xi(X) \right] \right| \\ &\leq \delta \sum_{\xi \neq 0} \left| \widehat{\Phi_{\mathbf{q}}}(\xi) \right|. \end{aligned}$$

It remains to bound the Fourier ℓ^1-mass of the test. Write

$$\Phi_{\mathbf{q}}(x) = \prod_{i=1}^{n} \varphi_i(x_i), \qquad \varphi_i(z) := \omega_m^{q_i(z)}.$$

Because both the uniform distribution and the character χ_ξ factor across coordinates, the Fourier coefficient also factorizes:

$$\widehat{\Phi_{\mathbf{q}}}(\xi) = \prod_{i=1}^{n} \mathbb{E}_{z \sim U_p} \left[\varphi_i(z) \omega_p^{-\xi_i z} \right].$$

Now suppose coordinate i is inactive. Then q_i is constant, so $\varphi_i(z) = c_i$ for some constant c_i of absolute value 1. The i-th factor in the Fourier coefficient is therefore

$$c_i \mathbb{E}_{z\sim U_p} \left[\omega_p^{-\xi_i z}\right].$$

If $\xi_i = 0$, this factor equals c_i. If $\xi_i \neq 0$, then

$$\mathbb{E}_{z\sim U_p} \left[\omega_p^{-\xi_i z}\right] = 0,$$

because a nontrivial character has average zero over $\mathbb{F}_p$. Therefore,

$$\widehat{\Phi_{\mathbf{q}}}(\xi) \neq 0$$

is possible only when

$$\xi_i = 0$$

at every inactive coordinate. There are w active coordinates. Each active coordinate ξ_i may take any of the p values in $\mathbb{F}_p$, while every inactive coordinate is forced to have $\xi_i = 0$. Hence there are at most p^w Fourier indices for which $\widehat{\Phi_{\mathbf{q}}}(\xi)$ can be nonzero. Since $|\Phi_{\mathbf{q}}(x)| = 1$ for every x, Parseval gives

$$\sum_{\xi\in\mathbb{F}_p^n} \left|\widehat{\Phi_{\mathbf{q}}}(\xi)\right|^2 = \mathbb{E}_U \left[|\Phi_{\mathbf{q}}(U)|^2\right] = 1.$$

Applying Cauchy–Schwarz only over the at most p^w nonzero Fourier coefficients,

$$\begin{aligned}\sum_{\xi\in\mathbb{F}_p^n} \left|\widehat{\Phi_{\mathbf{q}}}(\xi)\right| &\leq \sqrt{p^w} \left(\sum_{\xi\in\mathbb{F}_p^n} \left|\widehat{\Phi_{\mathbf{q}}}(\xi)\right|^2\right)^{1/2} \\ &= p^{w/2}.\end{aligned}$$

Substituting this into the previous estimate gives

$$\boxed{\left|\mathbb{E}_X \left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_U \left[\Phi_{\mathbf{q}}(U)\right]\right| \leq \delta p^{w/2}.}$$

This proves the lemma. □

5.2 The high-weight estimate

We now prove the second estimate used in the TRANSFER argument. Unlike the low-weight estimate, this bound uses the pseudorandomness guarantee of the incoming source Y.

Fix primes $p < r$ and a modulus $m \geq 2$, and define

$$\boxed{\kappa_{r,m} := \sqrt{1 - \frac{4\sin^2(\pi/m)}{r^2}}.}$$

Lemma 5.2 (High-weight estimate)**.** *Let $\sigma : \mathbb{F}_r \to \mathbb{F}_p$ be a map that reaches every element of $\mathbb{F}_p$. Suppose Y is a distribution on $\mathbb{F}_r^n$ satisfying*

$$\mathrm{Err}_Y\big(\mathsf{Local}_n(r,m)\big) \le \eta,$$

let B be any independent distribution on $\mathbb{F}_p^n$, and define

$$X := \sigma(Y) + B.$$

Then, for every $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p,m)$ of weight w,

$$\boxed{\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \le \eta + 2\kappa_{r,m}^{w},}$$

where $U_p \sim U_p^n$.

Proof. We begin with a one-coordinate cancellation bound. Let ν be any distribution on $\mathbb{F}_p$ satisfying

$$\nu(a) \ge \frac{1}{r} \qquad \text{for every } a \in \mathbb{F}_p,$$

and let

$$q : \mathbb{F}_p \to \mathbb{Z}_m$$

be nonconstant. Let X_0, X_0' be independent samples from ν, and define

$$W := \omega_m^{q(X_0)}, \qquad W' := \omega_m^{q(X_0')}.$$

Since $|W| = |W'| = 1$,

$$\begin{aligned}\mathbb{E}_{X_0,X_0'\sim\nu}\left[|W-W'|^2\right] &= 2 - 2\,\mathrm{Re}\,\mathbb{E}_{X_0,X_0'\sim\nu}\left[W\overline{W'}\right]\\ &= 2 - 2\left|\mathbb{E}_{X_0\sim\nu}\left[W\right]\right|^2.\end{aligned}$$

Therefore

$$1 - \left|\mathbb{E}_{X_0\sim\nu}\left[\omega_m^{q(X_0)}\right]\right|^2 = \frac{1}{2}\mathbb{E}_{X_0,X_0'\sim\nu}\left[|W-W'|^2\right].$$

Since q is nonconstant, there exist $a, b \in \mathbb{F}_p$ such that

$$q(a) \ne q(b).$$

Two distinct m-th roots of unity are at distance at least

$$2\sin\frac{\pi}{m}.$$

Keeping only the two ordered events

$$(X_0, X_0') = (a,b) \qquad \text{and} \qquad (X_0, X_0') = (b,a)$$

in the nonnegative expectation above gives

$$\begin{aligned}1 - \left|\mathbb{E}_{X_0\sim\nu}\left[\omega_m^{q(X_0)}\right]\right|^2 &\ge \nu(a)\nu(b)\left|\omega_m^{q(a)} - \omega_m^{q(b)}\right|^2\\ &\ge \frac{4\sin^2(\pi/m)}{r^2}.\end{aligned}$$

Hence

$$\boxed{\left|\mathbb{E}_{X_0\sim\nu}\left[\omega_m^{q(X_0)}\right]\right| \le \kappa_{r,m}.}$$

In particular,

$$0 < \kappa_{r,m} < 1.$$

We now apply this one-coordinate bound to the full local phase. Fix

$$\Phi_{\mathbf{q}}(x) = \omega_m^{\sum_{i=1}^n q_i(x_i)} \in \mathsf{Local}_n(p,m),$$

and let $w = \mathrm{wt}(\mathbf{q})$ be its number of active coordinates. Condition first on a fixed value

$$B = \beta.$$

The function

$$y \longmapsto \Phi_{\mathbf{q}}\big(\sigma(y)+\beta\big)$$

belongs to $\mathsf{Local}_n(r,m)$. Therefore the hypothesis on Y gives

$$\left|\mathbb{E}_Y\left[\Phi_{\mathbf{q}}\big(\sigma(Y)+\beta\big)\right] - \mathbb{E}_{U_r}\left[\Phi_{\mathbf{q}}\big(\sigma(U_r)+\beta\big)\right]\right| \le \eta,$$

where

$$U_r \sim U_r^n.$$

We next bound the second expectation. Since the coordinates of U_r are independent,

$$\mathbb{E}_{U_r}\left[\Phi_{\mathbf{q}}\big(\sigma(U_r)+\beta\big)\right] = \prod_{i=1}^n \mathbb{E}_{z\sim U_r}\left[\omega_m^{q_i(\sigma(z)+\beta_i)}\right].$$

Consider an active coordinate i. Since q_i is nonconstant and σ reaches every element of $\mathbb{F}_p$, the map

$$z \longmapsto q_i(\sigma(z)+\beta_i)$$

is nonconstant. Moreover, the random variable

$$\sigma(z)+\beta_i, \qquad z\sim U_r,$$

takes every value of $\mathbb{F}_p$ with probability at least $1/r$. Therefore the one-coordinate bound above gives

$$\left|\mathbb{E}_{z\sim U_r}\left[\omega_m^{q_i(\sigma(z)+\beta_i)}\right]\right| \le \kappa_{r,m}.$$

An inactive coordinate contributes a factor of absolute value 1. Hence

$$\left|\mathbb{E}_{U_r}\left[\Phi_{\mathbf{q}}\big(\sigma(U_r)+\beta\big)\right]\right| \le \kappa_{r,m}^w.$$

Combining this with the pseudorandomness guarantee for Y,

$$\left|\mathbb{E}_Y\left[\Phi_{\mathbf{q}}\big(\sigma(Y)+\beta\big)\right]\right| \le \eta + \kappa_{r,m}^w.$$

Averaging over the random vector B and using

$$X = \sigma(Y) + B$$

gives

$$\begin{aligned}\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right]\right| &= \left|\mathbb{E}_B\left[\mathbb{E}_Y\left[\Phi_{\mathbf{q}}\big(\sigma(Y)+B\big)\right]\right]\right| \\ &\le \mathbb{E}_B\left[\left|\mathbb{E}_Y\left[\Phi_{\mathbf{q}}\big(\sigma(Y)+B\big)\right]\right|\right] \\ &\le \eta + \kappa_{r,m}^{w}.\end{aligned}$$

It remains to bound the expectation under the target uniform distribution. Since $U_p \sim U_p^n$ has independent coordinates,

$$\mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right] = \prod_{i=1}^{n} \mathbb{E}_{z\sim U_p}\left[\omega_m^{q_i(z)}\right].$$

For an active coordinate, the uniform distribution on $\mathbb{F}_p$ assigns probability

$$\frac{1}{p} \ge \frac{1}{r}$$

to every value. Therefore the same one-coordinate bound gives

$$\left|\mathbb{E}_{z\sim U_p}\left[\omega_m^{q_i(z)}\right]\right| \le \kappa_{r,m}.$$

Thus

$$\left|\mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \le \kappa_{r,m}^{w}.$$

Finally, by the triangle inequality,

$$\begin{aligned}\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| &\le \left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right]\right| + \left|\mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \\ &\le \eta + 2\kappa_{r,m}^{w}.\end{aligned}$$

This proves the lemma. □

5.3 The TRANSFER theorem

The two estimates needed for $\mathrm{TRANSFER}$ have now been proved. For a target test of weight w, Lemma 5.1 gives

$$\text{error} \le \delta p^{w/2},$$

while Lemma 5.2 gives

$$\text{error} \le \eta + 2\kappa_{r,m}^{w},$$

where η is the error of the incoming source.

The first bound increases with w, while the second decreases with w. The proof of $\mathrm{TRANSFER}$ therefore only needs to choose a threshold at which we switch from the first estimate to the second.

Theorem 5.3 (TRANSFER)**.** *Let $p < r$ be primes, let $m \ge 2$, and let $0 < \varepsilon \le 1$. Fix a map $\sigma : \mathbb{F}_r \to \mathbb{F}_p$ that reaches every element of $\mathbb{F}_p$, and set*

$$\kappa = \kappa_{r,m}, \qquad W = \left\lceil \frac{\log(4/\varepsilon)}{\log(1/\kappa)} \right\rceil, \qquad \delta = \frac{\varepsilon}{2p^{W/2}}.$$

Suppose Y is a distribution on $\mathbb{F}_r^n$ such that

$$\mathrm{Err}_Y(\mathsf{Local}_n(r,m)) \le \frac{\varepsilon}{2},$$

and let B be an independent δ-biased distribution on $\mathbb{F}_p^n$. Define

$$X := \sigma(Y) + B.$$

Then

$$\boxed{\mathrm{Err}_X(\mathsf{Local}_n(p,m)) \le \varepsilon.}$$

If Y can be sampled with seed length

$$O_{r,m}\left(\log n + \log \frac{1}{\varepsilon}\right),$$

then X can be sampled with seed length

$$O_{p,r,m}\left(\log n + \log \frac{1}{\varepsilon}\right).$$

Proof. Fix a target test

$$\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p,m)$$

and let

$$w = \mathrm{wt}(\mathbf{q})$$

be its number of active coordinates. By Lemma 5.1, applied with $Z = \sigma(Y)$,

$$\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \le \delta p^{w/2},$$

where $U_p \sim U_p^n$. By Lemma 5.2, applied with

$$\eta = \frac{\varepsilon}{2},$$

we also have

$$\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \le \frac{\varepsilon}{2} + 2\kappa^w.$$

Thus every target test satisfies both estimates

$$\boxed{\left|\mathbb{E}_X\left[\Phi_{\mathbf{q}}(X)\right] - \mathbb{E}_{U_p}\left[\Phi_{\mathbf{q}}(U_p)\right]\right| \le \min\left\{\delta p^{w/2}, \frac{\varepsilon}{2} + 2\kappa^w\right\}.}$$

We now choose the estimate according to the value of w. If

$$w \le W,$$

then the low-weight estimate gives

$$\delta p^{w/2} \le \delta p^{W/2} = \frac{\varepsilon}{2} \le \varepsilon.$$

If

$$w > W,$$

then $0 < \kappa < 1$, and by the definition of W,

$$\kappa^w \le \kappa^W \le \frac{\varepsilon}{4}.$$

Hence the high-weight estimate gives

$$\frac{\varepsilon}{2} + 2\kappa^w \le \frac{\varepsilon}{2} + 2 \cdot \frac{\varepsilon}{4} = \varepsilon.$$

Therefore every $\Phi_{\mathbf{q}} \in \mathsf{Local}_n(p, m)$ has distinguishing error at most ε, and so

$$\mathrm{Err}_X(\mathsf{Local}_n(p, m)) \le \varepsilon.$$

It remains to check the seed length. Since

$$0 < \kappa < 1$$

is a constant depending only on r and m,

$$W = O_{r,m}\left(\log \frac{1}{\varepsilon}\right).$$

Therefore

$$\begin{aligned} \log \frac{1}{\delta} &= \log \frac{2}{\varepsilon} + \frac{W}{2} \log p \\ &= O_{p,r,m}\left(\log \frac{1}{\varepsilon}\right). \end{aligned}$$

An explicit δ-biased vector over $\mathbb{F}_p^n$ can therefore be sampled using

$$O_{p,r,m}\left(\log n + \log \frac{1}{\varepsilon}\right)$$

seed bits. Adding these bits to the seed used to sample Y gives the claimed seed length for X. $\square$

6 Assembling the generator for an arbitrary modulus

The two construction moves are now complete. This section follows the route from Section 2: start with the trivial modulus 1, insert the prime factors of M from largest to smallest, and finish over the Boolean alphabet.

Call a state (p, m) *solvable* if, for every n and every $0 < \varepsilon \le 1$, there is an explicit source X on $\mathbb{F}_p^n$ satisfying

$$\mathrm{Err}_X(\mathsf{Local}_n(p, m)) \le \varepsilon$$

with seed length

$$O_{p,m}\left(\log n + \log \frac{1}{\varepsilon}\right).$$

The state $(p, 1)$ is solvable for every prime p: the test class contains only the constant function 1, so the constant-zero source has error 0 and uses no seed.

Theorems 4.5 and 5.3 give two rules:

$$\begin{aligned} &(p, m) \text{ solvable} \implies (p, pm) \text{ solvable}, \\ &(r, m) \text{ solvable} \implies (p, m) \text{ solvable} \qquad (p < r). \end{aligned}$$

For the first rule, a target error ε requires the previous state only at accuracy

$$\left(\frac{\varepsilon}{4}\right)^{2^{p-1}},$$

whose reciprocal has logarithm $O_p(\log(1/\varepsilon))$. For the second rule, a target error ε requires the previous state at accuracy $\varepsilon/2$, and the added small-bias vector also has logarithmic seed length. Thus both rules preserve solvability.

Theorem 6.1 (Assembly over the prime factors)**.** *For every fixed integer $M \geq 2$, every n, and every $0 < \varepsilon \leq 1$, there is an explicit source G on $\mathbb{F}_2^n = \{0, 1\}^n$ such that*

$$\boxed{\mathrm{Err}_G(\mathsf{Local}_n(2, M)) \leq \varepsilon}$$

and whose seed length is

$$O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

Proof. Write the prime factorization of M, with multiplicity, as

$$M = p_1 p_2 \cdots p_k, \qquad p_1 \leq p_2 \leq \cdots \leq p_k.$$

Begin with the solvable state $(p_k, 1)$ and apply LIFT_{p_k}, obtaining (p_k, p_k).

Now process $j = k-1, k-2, \ldots, 1$. Suppose the already inserted suffix of the modulus is

$$m_j = p_{j+1} \cdots p_k$$

and the current state is (p_{j+1}, m_j). If $p_j < p_{j+1}$, first apply $\mathrm{TRANSFER}_{p_{j+1} \to p_j}$ to obtain (p_j, m_j). If $p_j = p_{j+1}$, this transfer is unnecessary because the alphabet already matches. In either case, apply LIFT_{p_j} to reach

$$(p_j, p_j m_j).$$

After all factors have been inserted, the state is (p_1, M). If $p_1 > 2$, one final application of $\mathrm{TRANSFER}_{p_1 \to 2}$ reaches $(2, M)$; if $p_1 = 2$, no final transfer is needed.

The number of moves depends only on the fixed modulus M, and each move preserves solvability. Hence the final state has the stated error and seed length. $\square$

7 From phase estimates to modular-sum distributions

Theorem 6.1 gives a Boolean source whose expectation is close to uniform for every modulus-M local phase. The main theorem is stated in terms of statistical distance between residue distributions. This final section converts the phase estimates into that distributional statement. The argument below is proven in [LRTV09] and is proved here again formally for completeness.

7.1 The phase guarantee supplied by the construction

Theorem 7.1 (Phase-error generator). *For every fixed $M \geq 2$, every n, and every $0 < \gamma \leq 1$, there is an explicit source G on $\{0,1\}^n$ with seed length $O_M(\log n + \log(1/\gamma))$ such that, for every $a \in \mathbb{Z}_M^n$,*

$$\boxed{\left|\mathbb{E}\omega_M^{\sum_i a_i G_i} - \mathbb{E}\omega_M^{\sum_i a_i U_i}\right| \leq \gamma,}$$

where U is uniform on $\{0,1\}^n$.

Proof. Apply Theorem 6.1 with error γ. For every $a \in \mathbb{Z}_M^n$, define $q_i(0) = 0$ and $q_i(1) = a_i$. Then

$$\omega_M^{\sum_i a_i x_i} = \omega_M^{\sum_i q_i(x_i)}$$

is a member of $\mathsf{Local}_n(2, M)$, so the displayed estimate is exactly the error guarantee from Theorem 6.1. □

7.2 Fourier inversion on $\mathbb{Z}_M$

Lemma 7.2 (From phase error to statistical distance). *Let X be a bit distribution such that, for every $b \in \mathbb{Z}_M^n$,*

$$\left|\mathbb{E}\omega_M^{\sum_i b_i X_i} - \mathbb{E}\omega_M^{\sum_i b_i U_i}\right| \leq \gamma.$$

Then, for every $a \in \mathbb{Z}_M^n$,

$$\boxed{\operatorname{dist}\left(S_a^{(M)}(X), S_a^{(M)}(U)\right) \leq \frac{M-1}{2}\gamma.}$$

Proof. Fix a, and let P and Q be the probability mass functions of $S_a^{(M)}(X)$ and $S_a^{(M)}(U)$ on $\mathbb{Z}_M$. For $t \in \mathbb{Z}_M$, define

$$\widehat{P}(t) = \sum_{z \in \mathbb{Z}_M} P(z)\omega_M^{tz}, \qquad \widehat{Q}(t) = \sum_{z \in \mathbb{Z}_M} Q(z)\omega_M^{tz}.$$

Then

$$\widehat{P}(t) = \mathbb{E}\omega_M^{tS_a^{(M)}(X)} = \mathbb{E}\omega_M^{\sum_i (ta_i)X_i},$$

and similarly for Q. At $t = 0$ both coefficients equal 1. For $t \neq 0$, the vector ta is another member of $\mathbb{Z}_M^n$, so the hypothesis gives

$$|\widehat{P}(t) - \widehat{Q}(t)| \leq \gamma.$$

Fourier inversion on $\mathbb{Z}_M$ gives, for every residue z,

$$P(z) - Q(z) = \frac{1}{M} \sum_{t \in \mathbb{Z}_M} (\widehat{P}(t) - \widehat{Q}(t)) \omega_M^{-tz}.$$

The term $t = 0$ is zero, and the triangle inequality therefore gives

$$|P(z) - Q(z)| \leq \frac{1}{M} \sum_{t \neq 0} |\widehat{P}(t) - \widehat{Q}(t)| \leq \frac{M-1}{M} \gamma.$$

Summing this bound over the M residues,

$$\operatorname{dist}(P, Q) = \frac{1}{2} \sum_{z} |P(z) - Q(z)| \leq \frac{M-1}{2} \gamma.$$

$\square$

Theorem 7.3 (Main theorem). *For every fixed integer $M \geq 2$, every $n \geq 1$, and every $0 < \varepsilon \leq \frac{1}{2}$, there is an explicit generator*

$$G_{M,n,\varepsilon} : \{0,1\}^s \longrightarrow \{0,1\}^n, \qquad s = O_M\left(\log n + \log \frac{1}{\varepsilon}\right),$$

such that, for every $a \in \mathbb{Z}_M^n$,

$$\operatorname{dist}\big(S_a^{(M)}(G_{M,n,\varepsilon}), S_a^{(M)}(U_n)\big) \leq \varepsilon.$$

Proof. Set

$$\gamma = \frac{2\varepsilon}{M-1}.$$

Since $0 < \varepsilon \leq 1/2$ and $M \geq 2$, we have $0 < \gamma \leq 1$. Let G be the source from Theorem 7.1 for this value of γ. Its seed length is

$$O_M\left(\log n + \log \frac{1}{\gamma}\right) = O_M\left(\log n + \log \frac{1}{\varepsilon}\right).$$

Lemma 7.2 gives, for every $a \in \mathbb{Z}_M^n$,

$$\operatorname{dist}\big(S_a^{(M)}(G), S_a^{(M)}(U_n)\big) \leq \frac{M-1}{2} \gamma = \varepsilon.$$

$\square$

7.3 Explicitness and optimality

The construction uses a number of LIFT and TRANSFER steps depending only on M. Every alphabet map is a fixed lookup table, every added random vector comes from an explicit small-bias construction over a fixed prime field, and all operations are coordinatewise. Hence the output is computable in time polynomial in n.

The dependence on both n and ε is optimal up to the constant depending on M.

Proposition 7.4 (Reduction to a prime divisor)**.** *Let q be a prime dividing M. Every bit distribution that ε-fools Boolean linear sums modulo M also ε-fools Boolean linear sums modulo q.*

Proof. Given $b \in \mathbb{Z}_q^n$, define

$$a_i = \frac{M}{q} b_i \pmod{M}.$$

For every Boolean input x,

$$S_a^{(M)}(x) = \frac{M}{q} S_b^{(q)}(x) \pmod{M}.$$

The map

$$\iota : \mathbb{Z}_q \longrightarrow \mathbb{Z}_M, \qquad \iota(z) = \frac{M}{q} z$$

is one-to-one. Applying a one-to-one map to two distributions preserves their statistical distance. Therefore the distance between the two modulo-q sums equals the distance between the corresponding modulo-M sums, which is at most ε by hypothesis. $\square$

Consequently, any generator for modulus M also fools linear sums modulo a fixed prime divisor of M. The standard lower bound for that prime-modulus problem is

$$\Omega\left(\log n + \log \frac{1}{\varepsilon}\right)$$

[AGHP92], [LRTV09]. Hence the seed length in Theorem 7.3 is optimal up to a constant depending on M.

8 References

[AGHP92] N. Alon, O. Goldreich, J. Håstad, and R. Peralta, "Simple Constructions of Almost k-wise Independent Random Variables," *Random Structures & Algorithms*, 3(3):289–304, 1992.

[AMN98] Y. Azar, R. Motwani, and J. Naor, "Approximating Probability Distributions Using Small Sample Spaces," *Combinatorica*, 18(2):151–171, 1998. Supplies explicit small-bias distributions over non-binary alphabets, used here for every prime alphabet $p > 2$.

[BV10] A. Bogdanov and E. Viola, "Pseudorandom Bits for Polynomials," *SIAM Journal on Computing*, 39(6):2464–2486, 2010. Preliminary version in *FOCS 2007*.

[GKM18] P. Gopalan, D. M. Kane, and R. Meka, "Pseudorandomness via the Discrete Fourier Transform," *SIAM Journal on Computing*, 47(6):2451–2487, 2018. Preliminary version in *FOCS 2015*.

[LRTV09] S. Lovett, O. Reingold, L. Trevisan, and S. Vadhan, "Pseudorandom Bit Generators That Fool Modular Sums," *RANDOM 2009*. Their optimal-order construction covers every fixed prime-power modulus; the present paper imports nothing from it.

[Lov09] S. Lovett, "Unconditional Pseudorandom Generators for Low-Degree Polynomials," *Theory of Computing*, 5:69–82, 2009. Preliminary version in *STOC 2008*.

[LV25] C. H. Lee and E. Viola, "Pseudorandom Bits for Non-Commutative Programs," *40th Computational Complexity Conference (CCC 2025)*, 9:1–9:22.

[MZ09] R. Meka and D. Zuckerman, "Small-Bias Spaces for Group Products," *RANDOM 2009*.

[NN93] J. Naor and M. Naor, "Small-Bias Probability Spaces: Efficient Constructions and Applications," *SIAM Journal on Computing*, 22(4):838–856, 1993. Explicit small-bias distributions are the only pseudorandomness primitive imported by the upper bound.

[Smo87] R. Smolensky, "Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity," *19th Annual ACM Symposium on Theory of Computing (STOC 1987)*, 77–82.

[Vio09] E. Viola, "The Sum of d Small-Bias Generators Fools Polynomials of Degree d," *Computational Complexity*, 18(2):209–217, 2009. Preliminary version in *CCC 2008*.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il